1. Stable Learning

The operation of any Relativistic system, in which teaching can be construed as some form of control or regulation of a learning process, takes place within certain assumed constraints upon the subject and the task he is performing[2]. The assumptions are minimal and appeared in Chapter 1 Section 2.8 as the "first group" of assumptions. An SST controller must satisfy the implied conditions; like any other regulator. But, being a very rigid type of device, it must satisfy others as well.

1.1. The subject is able to interpret task situations as problems that are to be solved.

1.2. Any problem can be characterised as the need to bring about or satisfy a relation. This might be a very concrete relation (for example, a stimulus response relation; the relation between alphabetic characters and keyboard positions in typewriting) or it might be abstract, as it usually is in academic skills (a relation between kings and countries; the relation between models for hydraulic flow and heat flow).

1.3. The subject has a goal which is either to solve problems or to learn how to solve them. Although it is true that the goal is achieved when certain relations are brought about and maintained (against disturbances that periodically pose problems), the goal itself is not just a stable condition; it is a purpose or intention entertained by the subject. In other words, unlike the run of literature on system theory, a relativistic theory not only stipulates that the external observer conceives the subject as though he had a goal but asserts that the subject does have a goal i.e. that he subscribes to an experimental contract in the sense of Chapter 1, Section 1.5.

---

[2] Control may be taken as liberally as desired. Some term like "persuading" or "advising" or "catalysing" might easily replace "controlling". Saying, in this sense, that teaching is the control of learning is not intended to suggest an authoritarian (or, for that matter, a permissive) approach; the phrase itself is neutral on such issues though it later appears that a specially authoritarian approach to teaching is not recommended.

1.4. The subject's responses are goal directed or purposive; they constitute solutions to problems or (in some cases) they index the description of solutions. The problems that are solved may be divided into two categories; namely the application of problem solving procedures that exist and the construction of procedures, if none exist. Alternatively, the two categories may be called "problem solving" unqualified and "learning to solve problems".

1.5. It is emphasised that this distinction (as stated) is quite arbitrary insofar as it is made by an external observer and not (usually) by the student. Elshout and Elshout (1969) stress, in their discussion of solving the creative problems posed by Guilford's "apparatus test", that there is no fundamental distinction between the act of solving problems and the "higher level" problem solving that goes on when a problem solving procedure or strategy is constructed. This point is accepted in principle. Nevertheless, a distinction is made on the grounds of convenience and in order to partition SST operations.

1.6. The contents of subsections 1.1, 1.3, and 1.4 make sense only within the framework of a normative model (Chapter 1, Fig. 2) and given an object language used by the subject (as one participant) in which the participants can describe and prescribe problems, goals and so on. However, the extent to which language usage is obtrusive varies a great deal (permitting the extreme cases and the intermediary cases of Chapter 1, Section 1.5). So, for example, it is quite possible to talk of a tracking task SST system as though in stimulus/response terminology. But, as noted by Gaines (1972), this kind of account is untenable if the task is complicated; for example, by a periodic modification of vehicle characteristics. The SST mode of "participant interaction" (Appendix A) bears witness to the same fact.

1.7. To ensure that there is sufficient variety to overtax the subject, the experimental contract must hold under conditions that would present the unaided subject with insuperable problems.

1.8. A participant other than the subject may cooperate with the subject insofar as he can partially solve the problems that are presented to the subject. If the subject learns to solve problems, as he must do under the conditions prescribed in the (last) subsection 1.7, then the other participant must learn how to cooperate in order to maintain steady state conditions. On the one hand, this learning might be independent, i.e. the other participant regards

the subject as an *it* which behaves in a variable manner. On the other hand the learning might rest upon a dialogue between the participants in which problem solving methods are described (the CET of Chapter 1, Section 1.5).

2. Specific Additional Constraints on the SST System

The operation of the particular SST system considered in Chapter 1, Section 2 calls for more rigid and generally less plausible assumptions (the "second group" of Chapter 1, Section 2.8). These are summarised below, as applying to the subject and to the task. It will be evident that rather few task situations satisfy these restrictions; those that do so are called "structured skills" in the supporting literature (Pask et al. 1964, 1965, 1966, 1969, 1970). Fortunately, it happens that most common perceptual motor skills are (or can be instructed as) "structured skills". This is also true of small segments of educationally significant material (for example, people may learn to deal with each of the "topic relations", introduced in Chapter 4, as though they were learning a structured skill provided that the "topic relation" is isolated). The possibility of imposing such a structure upon otherwise amorphous items also accounts for the apparently tractable character of "drill and practice" instruction. Though there can be no doubt that structured perceptual motor skills are easily and effectively inculcated and though vocabulary lists must occasionally be learned (if so, then it is justifiable to impose an appropriately structured context in which to instruct them) the general utility of indoctrinating pupils in respect of isolated topic relations or of using the "drill and practice" method is very questionable. The results discussed later suggest that there is little merit in such dissections and that the corresponding techniques have only limited and local value in education.

The assumptions in question are as follows.

2.1. It must be possible to specify a class of operations, called simplifying operations, that map the task relations onto classes of relations or properties that, in mathematical terms, have either a lower cylindrance (Ashby 1964) or relate fewer properties or have fewer members; alternatively it must be possible to furnish cueing information that specifies the values of some coordinates of the task relation; either directly or by complementation (excluding solutions). Finally, a simplification must be repeatable to yield nested classes of more or less simplified forms of the task relation until, at the last stage, the relation is fully specified.

2.2. If such an operation can be performed uniquely to provide a linearly ordered nesting (Fig. 1) then the classes in question are indexed by a numerical variable η and the simplification scheme is unidimensional; if the nesting scheme is partially ordered with m branches (Fig. 1) then the simplification scheme is indexed by a vector and $\eta = \langle \eta_1 \ldots \eta_m \rangle$, with components η_i; $i = 1, \ldots m$. In the former case an SST system may (given the caveats noted below) be stabilised by a unidimensional controller and, in the latter case, by a multi-dimensional controller having m sub-controllers responsible for adjusting the η_i to maintain the performance indices ρ_i near to the fiducial values ξ_i.

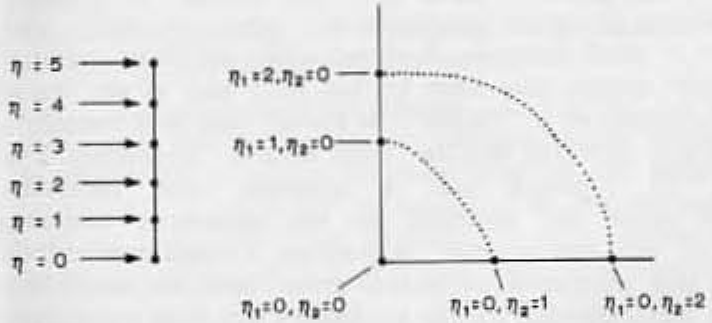

Fig. 2.1. A very simple image of the conditions needed for task simplification. Subtasks, indexed by η may be conceived as located at a distance d from the main task (at origin $\eta = 0$ or $\eta_1 = 0, \eta_2 = 0$). It is required for any linear structure that $d(\eta+1, 0) > d(\eta, 0)$. As an additional requirement, for the branching structure $d(\eta_1 + 1, \eta_2 + 1, 0) \geqslant d(\eta_1 + 1, \eta, 0)$ and that $d(\eta_1+1, \eta_2+1, 0) \geqslant d(\eta_1+1, \eta_2, 0)$.

2.3. Consider a linearly indexed simplification scheme. Let "Class η_0" stand for the class indexed by a value η_0 of η. Let "ρ (Class η_0)" stand for the value of performance index computed over an arbitrary sequence of problems posed under the task proper to Class η_0 when such problems are first encountered (thus ρ (Class η_0) is not defined until problems of this kind have been presented). Finally, let $\bar{\rho}$ (Class η_0) represent the (possibly variable) value of ρ for all subsequent presentations as part of an SST interaction. We require

$$\rho\,(\text{Class } \eta) > \rho\,(\text{Class } \eta + 1) \text{ and}$$
$$\bar{\rho}\,(\text{Class } \eta) \geqslant \bar{\rho}\,(\text{Class } \eta + 1)$$

for all values of η.

2.4 For an m fold indexing scheme each η_i and ρ_i must satisfy

$$\rho_i\,(\text{Class } \eta_i) > \rho_i\,(\text{Class } \eta_i + 1)$$
$$\bar{\rho}_i\,(\text{Class } \eta_i) \geqslant \bar{\rho}_i\,(\text{Class } \eta_i + 1)$$

where only one component of η, the i-th, is varied at once. In addition, any joint variation (written $\eta_i + 1, \eta_j + 1$) is more difficult, initially, than a one step variation. To express this requirement, it is convenient to employ certain abbreviations, namely

$$\text{Mean Value}\,[\bar{\rho}_i \cdots \bar{\rho}_j\,(\text{Class } \eta_i \cdots \eta_j)] \quad = \hat{\rho}_{ij}\,\eta$$
$$\text{Mean Value}\,[\bar{\rho}_i \cdots \bar{\rho}_j\,(\text{Class } \eta_i + 1 \cdot \eta_j + 1)] = \hat{\rho}_{ij}\,\eta + 1$$

when the following conditions are to obtain for all variations (unitary or joint) of η.

$$\rho_{ij}\,\eta > \rho_{ij}\,\eta + 1$$
$$[(\hat{\rho}_{ij}\,\eta + 1) - (\hat{\rho}_{ij}\,\eta)] > [\bar{\rho}_i\,(\text{Class } \eta_i + 1) - (\bar{\rho}_i\,(\text{Class } \eta_i)]$$
$$[(\hat{\rho}_{ij} \quad + 1) - (\hat{\rho}_{ij}\,\eta)] > [\bar{\rho}_j\,(\text{Class } \eta_j + 1) - (\bar{\rho}_j\,(\text{Class } \eta_j)]$$

2.5 Let ρ^*, ρ^*_i, represent the asymptotic values of a performance index, that is, the value of ρ or ρ_i as time goes by or trials are performed (the time or trials being devoted to the same subtask). For a unidimensional SST system (with ε set as an arbitrary small number and with $(1 \geqslant \rho \geqslant 0)$) it is required that for $\eta = 0, \rho^* > 1 - \epsilon$ and that, for all η, if $\rho^* > 1 - \epsilon$ for η then $\rho^* > 1 - \epsilon$ for $\eta + 1$.

This condition guarantees learnability; the subject can learn to perform the simplest subtask and for any choice of η the increment *indexed* by "+1" is small enough to ensure that he can also learn eventually to perform the next most difficult subtask.

For a multidimensional SST system, this condition is required (jointly) of each variable pair (η_i, ρ_i) taken alone. It is not necessarily required of each joint variation. Thus, for each i, if

$\eta_i = 0$ then $\rho^*_i > 1 - \epsilon$ and if $\rho^*_i > 1 - \epsilon$ for η_i then $\rho^*_i > 1 - \epsilon$ for $\eta_i + 1$.

2.6. The bite behind these apparently innocent requirements is that (in order to satisfy them) the controller must have a model of how the subject solves problems. Failing that, the "simplification" obtained (for example) by decreasing η will not necessarily have the effect required in Subsections 2.4 and 2.5. Conversely, the "simplifying operators" of Subsections 2.1 and 2.2 are the controller's representation of how the subject does solve problems.

If this is not so, attempted cooperation may have the reverse effect; there are plenty of pathological cases where partially solving a problem according to a method that is not adopted by the subject actively interferes with the problem solving; one interesting and well documented example is described by Campbell (1968) in connection with code word problems.

The controller can secure the necessary correspondence between the subject's method and its own in two different ways; by learning in terms of a dialogue about problem solving methods, which methods the subject is using (and it may, on this basis, effect an idiosyncratically tailored simplification scheme). Something like this is done by the CET based situations discussed from Chapter 4 onwards. But simple SST systems do not have this capability (and even the "metasystems" discussed towards the end of Appendix A have limited capabilities in this respect). On the other hand, any controller even quite a simple one, can impose its simplification scheme (supposing it exists) upon the subject, thus straightjacketing the problem solving methods he can learn.

There is no objection to this if all methods are in some sense equivalent or if one class of methods is optimal. Otherwise, there is every objection to such a restrictive expedient. Several specific objections will be voiced in the sequel.

3. Types of Learning and Teaching Theory

Given the general stability of a learning/teaching system it is possible to construct two kinds of theory about its operation. One kind is a *macrotheory* (or, in a different jargon, a *molar* theory) and is expressed in terms of statistical variables. These variables estimate, either directly or indirectly, indices of the uncertainty experienced by the subject as he deals with distinct classes of problem; this degree of belief or doubt is augmented by a measure

of the retitude of the belief. The other kind of theory is a *microtheory* (alias a *molecular* theory). It deals in terms of mental events, linguistic transactions, and the like and purports to give an account of an underlying cognitive mechanism.

Both kinds of theory are developed in the sequel.

There is a sense in which a competent microtheory must predict macroscopic events. Later, it will be possible to make certain predictions. But it is also, and quite profoundly, true that the macrotheory of mechanism is only occasionally able to say what the subject's uncertainty (for example) is; in contrast to the much easier job of predicting when it will change and in what direction. Until the contrary is stated, it will be prudent to regard the macrotheory and the microtheory as separate entities.

4. A Macrotheory of Stable SST Interactions

One common feature of all the proficiency indices (of which several were suggested in Section 1) is that they estimate the student's uncertainty (call it I) about the solution of a class of problems. One common feature of all the difficulty indices η of Section 1.4 is that they estimate the structural uncertainty (call in I*) of a task situation and, as a consequence, the maximum uncertainty a student might experience at the instant in question. There is also a "most difficult" situation, when η is increased to the upper limit with an overall maximum uncertainty I^*_{max}.

Suppose (as part of a macrotheoretic orientation) that we are not concerned with the student's method of computation (any method will do) it is legitimate to regard the problem solving process in the student as a "Black Box" in Ashby's (1964) sense and to associate the "Black Box", at a given instant, with an uncertainty reduction or "information transfer" term, J, indicating his average ability, at this instant, to select amongst the possible solutions to a class of problems when randomly sampled problems are given. All of these uncertainties may be written vectorially, if required using standard multi-variate information theory, for example as it is discussed by Ashby (1969), Garner (1962), Stanisland (1966), McGill (1963), Attneave (1959) and Conant (1968).

An appropriate representation of learning in a stable SST system is a straightforward interpretation of Von Foerster's (1960) theory of self-organising systems. It is crucial that an organisation Z, (such as the redundancy function) is a function of two statistical variables of the type I, I*, for example, $Z = 1 - I/I^*$. A system is self-organising if and only if $dZ/dt > 0$ and this requirement is satisfied (over a region bounded between 0 and I^*_{max}) by an appropriately coupled adjustment of I and I*.

A stable adaptive teaching system is identified with a self-organising system by noting that $I = I^* - J$ and that (for fixed I*) adaptation, increasing J, leads to a change dJ/dt. It is further postulated, in Fig. 2, that the rate dI/dt of uncertainty reduction (learning) is a function of $I = I^* - J$ (loading) with a single maximum. Hence a steady state condition is achievable by adjusting I* and for a single variable system the maximum learning rate is obtained by choosing dI^*/dt so that $I = I^* - J = \gamma$. In general it is maintained that an adaptive learning-control system is stable if $dZ/dt > 0$ and it is operated at a (local) optimum if dI^*/dt and γ are chosen to maximise this rate.

If it happens that all reduction in uncertainty is due to an increase in correct certainty, then it can be shown (Cowan 1962) that all of the proficiency indices ρ we have employed are monotonically related to $-$ I. Hence, $J^* \equiv \eta$ (in any case), $-I \equiv \rho$ and γ (of Fig. 2) $\equiv \xi$ (Fig. 5 of Chapter 1).

Thus interpreted, ρ is a correct response frequency. The information or uncertainty index is computed on the assumption that the frequency estimates the correct response probability and consequently I is given by a function such as $I = \rho \log \rho + (1 - \rho) \log (1 - \rho)$. For most of the skills discussed in the literature this form is appropriate.

Whilst I* is held constant, this interpretation of the macrotheory corresponds to statistical learning theory. All the elegant

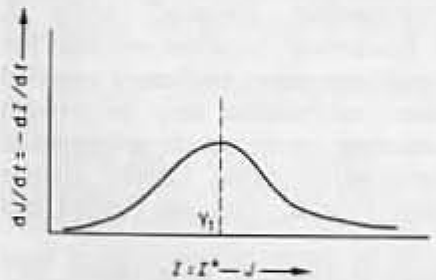

Fig. 2.2. Sketch showing form of uncertainty reduction rate (in one sense, "learning rate") as function of uncertainty level.

concepts due to Estes and many others (see, for example, the collected papers in Luce, Bush and Gallanter 1963a, 1963b) may be rewritten in terms of "I functions" and the like where the "I functions" stands for an external observer's uncertainty/information (given a model) about the value of certain variates. The exercise is open to Von Foerster's criticism (mentioned in Section 1.5.1.2) on structural, not mathematical grounds. If the organism/environment "boxes" are finite state machines (as they are if I* can be kept fixed) then all learning curves plotted in I (or ρ) are predeterminable and the only remaining issue is which kind of adaptation does take place.

The present macrotheory departs from statistical learning theory in two respects. In the first place I* is not kept constant and the criterial variate, the organisation, Z, is a function of I* as well as I. Reduction in I alone is adaptation (albeit goal directed in form) but it does not count as learning which, within this theory, entails a change in what may occupy the subject's attention (a change in I*, of course) which is correlated through the other participant (albeit an SST controller or an adaptive teaching device) with changes in I.

The other distinction is closely related. As announced in the first paragraph, I is taken as an estimate of the subject's uncertainty about how to solve problems (recall the caution in Section 1.5.1.3) not the external observer's uncertainty about the currently exhibited problem set. The relative values of I and I* are maintained by the regulator i.e. the variables are relativistic quantities.

In view of these comments it is just as legitimate to estimate I, I*, directly as degrees of belief, rather than indirectly from correct response frequencies, and the like. Whenever possible, this is done and the control variables are written H, H*, rather than I, I*, to signify this fact. For example, in the CET code learning experiments, it was psychologically possible to obtain confidence estimates from which to compute a quantity (similar to H*) and to use a sequential guessing procedure as a psychologically feasible approximation to H. In general, for intellectual skills, H* is computed as $-\sum_i p_i \log p_i$ at each trial over sets of m alternatives $i = 1, \ldots, m$ for which the subject gives m degrees of belief p_i. In general also, it is possible to obtain an index, θ, of correct belief (we usually employ the simplest of all Schuford scoring functions; $\theta = 1 + \log p_{correct}$). Finally, the value of H^* is estimated by eliciting confidence estimates over problems that a subject can appreciate though he can not yet learn to solve them.

5. A Microtheory of Learning and Teaching

The following postulates sketch out a theory of learning and its (dual) theory of teaching tailored to fit a relativistic conversational system. The theory is presented properly in Chapters 5 and 6 on foundations laid down in Chapter 4; the sketch gives enough of the flavour of the ideas for the reader to make sense of the experimental work considered in Chapter 3. The main tenets of the theory have been confirmed in the laboratory. The postulates themselves also stand up quite well to the kind of linguistic analysis favoured by philosophers (i.e. the basic definitions give ordinary language meanings to such terms as "concept" and "memory") provided that certain misusages are noted at the outset. In particular, data storage (which surely exists) is not memory. As a matter of fact few people seriously believe it is; but a transposition of "memory", as used in computer jargon to refer to a storage location, is often regarded as harmless; in a relativistic system, the misnomer is also misleading. By the same token (as noted before) neither adaptation nor habituation (both of which occur as respectable processes in their own right) are of the same kind as learning; in a relativistic system, learning must be a form of evolution (in a literal not a figurative sense, as stressed by reference to Von Foerster's 1971 reductio ad absurdum). Finally a "concept" must be some kind of process. Once again, this point is generally conceded but there is a very pervasive habit of loosely identifying concepts with the classes or categories which (amongst other things) they may construct or operate upon.

5.1. Whatever may be done (generally, to "solve a problem") is to bring about a relation. In the context of intellectual problem solving, it is apt to call a relation a *topic* (the perceptual motor relations between signals, response operations etc. brought about by exercising skills have already been exemplified). Connected collections of topic relations on which discourse between a pair of participants might dwell or which might occupy a participant's attention are called a *domain*; in general a *conversational* domain.

5.2. Problem solving is derivable from two equivalent processes, one is "making a model" (perhaps building an artifact) that, on

execution, brings about a relation; this kind of problem is posed by a command. The other is a linguistic process that explains a relation in the conversational language, indicating how it may be brought about (the problem is posed by a how or why question). Both processes bring about a topic relation when it does not currently exist. Other modes of problem solving are readily adumbrated as special instances of these two; for example, some part of a relation may hold already, when the relation is satisfied by less comprehensive operations. There are common occasions when the other participant in a conversation furnishes part of an explanation or specifies some of its "dimensions" (the corresponding problems are posed either by a qualification of the original why question or by asking a which or whether question).

5.3. A participating subject entertains an intention made explicit in the experimental contract to which he subscribes, as aiming for a certain class of *goals* or equivalently aiming to solve a certain class of problems.

5.4. As noted in Section 1.5 it is convenient, for descriptive purposes, to introduce an arbitrary distinction between levels of goal directed activity, which is a distinction identical with the control engineer's distinction between levels of control.

5.5. It is important to notice that these levels are not simply nestings of subroutines, as familiar constructs like an hierarchy of TOTE units (Miller, Gallanter and Pribram 1960) are ambiguous in this respect. For example, the original (1960) statement of a TOTE unit and its connections may be literally interpreted as designating a unit of control; in which case a TOTE hierarchy determines a level of control in the present sense (I have used this interpretation in many previous publications; for example Pask 1963, 1964, 1968a, 1968b, 1970b, 1971). On the other hand, a later publication (Miller and Chomsky 1963) gives a restricted and much less comprehensive interpretation to a TOTE unit, and consequently a TOTE hierarchy (which becomes a partition of an automaton, possibly represented as a subroutine of a serial program corresponding to the original automaton). The usage of this and closely allied constructs in both the psychological and system theoretic literature is variable, sometimes even careless, so it is necessary to insist that a very definite meaning is intended. It is symbolised in Fig. 3. Any problem solver in the lower box operates upon a domain of topic relations; to bring them about or

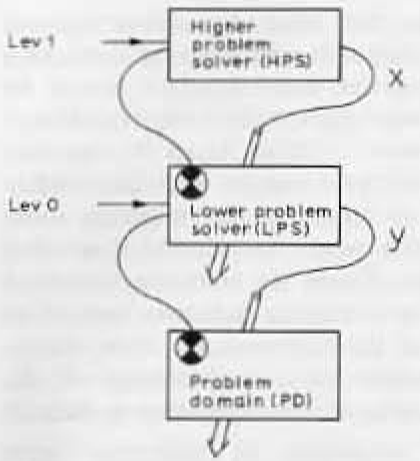

Fig. 2.3. Problem Solving and Learning.

to explain them, being in receipt of a description furnished either by itself (the ⊗ notation) or by a stretch of dialogue. Its operations are represented by the parametric or constructive arrow (the ⇓ notation). Any problem solver in the upper box operates in the same fashion upon a collection of problem solvers, to construct them de novo, or to reconstruct them, acting, once again, on the basis of a description. Either box may contain problem solvers of arbitrary complexity; for example, containing many subroutines. But, because of the arbitrary distinction of domain, the upper box is engaged in "problem solving of problem

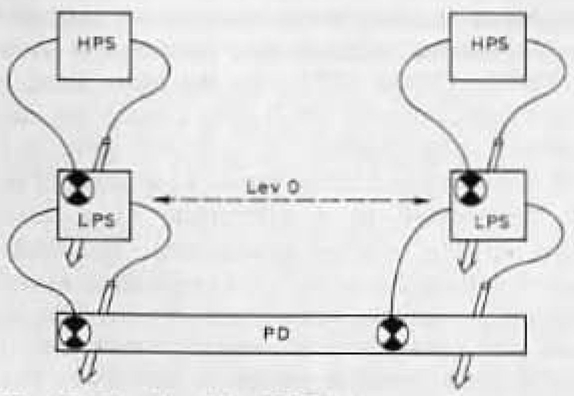

Fig. 2.4. Shorthand for SST System.

solving" or "learning". By the same token problem solving/learning in a participant can be controlled by another participant at two levels of control shown as Lev 0 and Lev 1 in Fig. 3. We take this arrangement as the least structure needed to characterise a participant and comment that the least interpretation placed upon it is a pair of reproductive systems (in the abstract, rather than the biological sense) together with couplings or connections at Lev 0 and Lev 1. The structure is by no means peculiar to human participants. For example the arrangement in Chapter 1, Fig. 6 has a shorthand representation in Fig. 4, and the full relativistic CET system is shown in shorthand as Fig. 5.

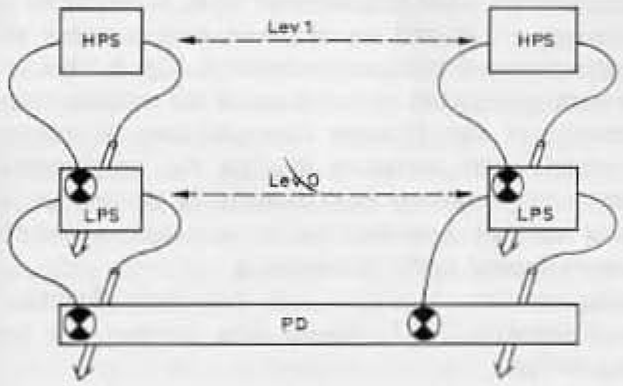

Fig. 2.5. Shorthand for CET System.

5.6. A problem solver is a goal directed procedure or a class of them. It is written (as a procedure to be executed) in the conversational object language. The term procedure could mean "serial program" but it is essential to include (Manna 1970) non-deterministic procedures (i.e. programs that have an execution tree rather than a line on which the locus of control occupies one point) and Fuzzy algorithms (Zadeh 1973) which is exactly what we mean by the word *heuristic*. Certain other concurrent procedures are to be examined in the sequel.

5.7. A *concept* of a topic relation (call it R_i) is defined as a procedure that brings about R_i; that is concept i $\triangleq$ reproduction (R_i).

This definition seems perverse because it does not tally with the usual notion of a concept as (something like) a "class" or "a description of a class" or "a stored description". But the definition is really commonsensical. Concepts are used to do things; in general, to reconstruct; that is, to reproduce; that is, to stabilise relations (Section 5.1); hence, amongst other things, classes. In the special cases of Section 5.2 concepts may recognise or select members of classes. Likewise, in any Lev 0 transaction, a concept explains a relation (conceivably how to generate a class like "orange rectangles") and, in the special cases of Section 5.2 it may answer questions about class membership. The description of a concept (for example in a Lev 1 transaction) contains at least one hypothesis.

5.8. A *memory* is defined as the reconstruction (alias the reproduction) of a concept. That is, a memory of R_i is Memory i $\triangleq$ reproduction (Concept i). It will be observed that a stable or realisable concept is depicted by the construct of Fig. 3 taken in isolation. Insofar as the permitted connections of the construct are realised (for example, in Fig. 5) some concepts may be jointly reconstructed, perhaps with variation due to the participants ostending different topic relations in a domain. If the matter is phrased in terms of abstract reproduction, the process in question is evolution. In psychological terms it is learning.

The participating entities in Fig. 5 are the finite function machines in Von Foerster's (1971) theory. The limiting case for non trivial learning is Fig. 4.

5.9. For any (conscious) participant we postulate that the upper and the lower control loops (labelled x and y in Fig. 3) remain active at a certain minimal rate. That is (lower loop) a participant is impelled to solve problems and (upper loop) he is built to learn. The macrotheoretic changes of uncertainty are due to this joint activity.

5.10. Within such a dynamic system, certain invariances necessarily exist. If any part of the system is isolated (as it would be, for example, if one participant learned in a particular way, exhibited in dialogue as a learning strategy) then it is necessarily true that a trapping condition will occur. In psychological terms this is a cognitive fixity whereby a plan, pattern or organisation, once established, becomes ingrained. It is manifest at all levels of mentation (consider, for example, the constraint imposed upon writing a paper by the happy or unhappy construction of a first paragraph or a notation scheme).

Festinger's (1957) cognitive dissonance is a special case bearing

upon an hypothesis (Section 5.7) in a concept; data refuting the currently exposed (trapping) hypothesis are rejected or perverted to maintain the hypothesis in force i.e. to reproduce it.

5.11. Certain types of rejection or incompatibility also necessarily exist; they are analogous to immune responses in physiology. Broadly (the matter is best discussed later), any reproducible procedure has a description of "other than itself" which it can recognise and certain "other than itself" descriptions are rejected as incompatible; conversely, only specific clusters of reproducible procedures can exist in the context of a given topic relation i.e. in a particular part of the conversational domain.

It is thus possible to stipulate, for any relativistic system, of two or more participants anchored on the same domain, that some classes of procedure are incompatible, and of any one participant that (whilst he remains in the same domain) certain otherwise possible procedures are incompatible with whatever class of procedures are embedded by cognitive fixity (Section 5.10). Particular interest is attached to procedures in the upper box of Fig. 3 as these determine how concepts are assembled i.e. what learning strategy is adopted. The theory predicts that certain classes of strategy will, in this sense, be exclusive; one distinction of this type is examined in the next chapter (more or less concurrent learning strategies, dubbed holist and serialist).

5.12. Consider the two participants in Fig. 5 exerting control upon one another. Call a condition of matching (not necessarily one to one) "agreement". It is possible to establish either Lev 0 or Lev 1 agreement or both.

If a participant explains a topic relation at Lev 0, this is evidence for a concept i; if the explanation is agreed, that is evidence for a concept equivalent to (not necessarily identical with) a concept entertained by the other participant. If the participant explains how he constructed and reconstructs this concept, at Lev 1, this is evidence for a memory; if the explanation is agreed, in the sense that it reproduces an equivalent concept in the other participant, that is evidence for an equivalent (not necessarily identical) memory. This condition is called *understanding* (in a given domain, by these participants, of a topic relation).